# SIMULTANEOUS GENOMIC DETECTION OF MULTIPLE ENTERIC BACTERIAL AND VIRAL PATHOGENS, INCLUDING SARS-CoV and RVFV

Shelley Payne
The University of Texas at Austin
Austin, TX 78712-0162

C.J. Peters and Shinji Makino
The University of Texas Medical Branch at Galveston
Galveston, TX 77555-1019

Kerry Oliver and Christy Weiss
Radix BioSolutions, Ltd.
Georgetown, TX 78626

Steve Kornguth*, Lenny Carruthers, and Robert Chin
Center for Strategic and Innovative Technologies, The University of Texas at Austin
Austin, TX 78759-5316

## ABSTRACT

A multiplexed screening system to detect pathogenicity islands (PI) of bacteria causing enteric disease and pathogenicity factors (PF) associated with the SARS-associated coronavirus (SARS-CoV) and Rift Valley Fever Virus (RVFV) has been developed. Pathogenic bacteria contain DNA sequences (PI) that code for proteins enhancing the ability of the bacteria to cause disease in the host, while viruses possess genomic sequences (PF) that are associated with their pathogenicity. By quickly screening for multiple pathogenicity island and factor sequences, end-users will have the capability to detect the first signs of an enteric or viral bioattack without requiring screening for a particular organism.

## INTRODUCTION

The presence of pathogens in the environment is usually recognized only after illnesses have resulted from infection by the pathogen or its toxic products. The time interval between exposure of a target to bacterial or viral threat agent and the development of clinical signs approximates between 36 hours to ten days. During this interval, the exposed subject may disseminate disease to a second generation of patients and exacerbate the management problem; however, antibiotics and antivirals can also be administered to ameliorate the development of disease. Toxins can cause clinical signs to appear within minutes to days but the dissemination of disease to a new population is very unlikely. The impact of these infectious agents on human and livestock targets would be greatly reduced if the threat agent is detected prior to or soon after exposure.

Pathogenic bacteria, viruses and fungi can be detected by a variety of antibody, metabolic, or DNA based assays; toxins are most readily detected by antibody based sensors or devices that measure the effect of the toxin on metabolism of cells. If performed appropriately, genomic based detection systems readily detect known threat agents with a high degree of certainty (very low to no false positives or negatives). This is because the pathogenic properties of an organism reside in the genome. For example, acquisition of the phage-borne pathogenicity island *slt* gene by nonpathogenic *E. coli* could lead to the

# Report Documentation Page

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| **15 NOV 2004** | **N/A** | **-** |

| 4. TITLE AND SUBTITLE | 5a. CONTRACT NUMBER |
|---|---|
| **Simultaneous Genomic Detection Of Multiple Enteric Bacterial And Viral Pathogens, Including SARS-CoV and RVFV** | 5b. GRANT NUMBER |
| | 5c. PROGRAM ELEMENT NUMBER |
| 6. AUTHOR(S) | 5d. PROJECT NUMBER |
| | 5e. TASK NUMBER |
| | 5f. WORK UNIT NUMBER |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **The University of Texas at Austin Austin, TX 78712-0162** | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. SPONSOR/MONITOR'S ACRONYM(S) |
| | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) |

12. DISTRIBUTION/AVAILABILITY STATEMENT
**Approved for public release, distribution unlimited**

13. SUPPLEMENTARY NOTES
**See also ADM001849, 2004 Scientific Conference on Chemical and Biological Defense Research. Held in Hunt Valley, Maryland on 15-17 November 2004 . , The original document contains color images.**

14. ABSTRACT

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT **unclassified** | b. ABSTRACT **unclassified** | c. THIS PAGE **unclassified** | **UU** | **8** | |

emergence of hemorrhagic colitis caused by a non-O157:H7 strain. Of equal concern is the possibility that the toxin gene, along with other essential virulence determinants, could be introduced into an *E. coli* or *Shigella* strain that does not match the known pathogenic serotypes or even into a bacterial strain that causes respiratory infections, thus delivering the toxin through a potentially more lethal route. For this reason, tools to rapidly identify specific genes that are essential for production of disease in the host are being developed. By screening for a large number of virulence determinants, many of which are found within pathogenicity islands, patterns of genes present in a potential pathogen can be identified. The set of virulence determinants present indicates the type of disease that could be produced. The presence of genes required for virulence is a more reliable indicator of pathogenicity than antigenic or metabolic characteristics that could vary independently of disease potential. This is because the antigenic properties of many pathogens are very similar to that of non-pathogenic organisms (e.g. bacillus anthracis and bacillus cereus) and so false positives are confounding.

With the emergence of annual outbreaks of specific pathogenic viruses (e.g., SARS, West Nile Virus), along with the danger of viral bioattack, it has become increasingly important to rapidly detect and identify these and other pathogenic viruses. With respect to the SARS-CoV, it has an enveloped positive-stranded RNA virus with a single stranded genome between 27 kb and 31.5 kb. The following are some of the structural genes found in this virus: S, E, M and N. The S protein is a major structural protein that mediates the fusion activity that is observed in coronavirus-infected cells and it plays an important role in virus attachment to the coronavirus-specific receptor. The M and E proteins are essential for efficient production of virus envelope and release (1). Additionally, the M protein plays a critical role in incorporation of S and N proteins into virus particles (2-5). With respect to RVFV, it too has several known genes, two of them being NSs and L (6, 7). All of these relatively conserved, structurally critical genes found in SARS-CoV and RVFV may serve as specific identifiers to the presence of a specific virus.

This report describes the preparation of synthetic DNA probes designed to detect amplicons of enteric pathogenic bacteria, the SARS-CoV, and the RVFV and their use with a multiplexed platform that has the capability to assay several hundred samples per hour. The platform is based upon the Luminex xMAP™ System, a multiplexed assay platform that combines high sample throughput (up to 600 samples/hour) with high information content per sample (up to 100 parameters tested simultaneously per sample). Additionally, this detection system will have hundreds of probes (multiplexed), so false positives and false negatives can be reduced and managed. The challenge of this multiplexed system is to develop appropriate probes that will read multiple amplicons in a mixture of materials with minimal interference and cross reactions between the selected probes or amplicons. An associated challenge is to develop probes that operate at similar melting temperatures so that a single protocol will be appropriate for assay conditions.

## MATERIALS & METHODS

### 1. PREPARATION OF AMPLICONS AND GENOMIC PROBES FOR ENTEROPATHOGENIC BACTERIA, THE N AND M GENES OF THE SARS-CoV AND THE NSs AND L GENES OF THE RVFV

Biotinylated amplicons for pathogenicity islands of *Shigella dysenteriae* and *S. flexneri* (*shuA, stx, ipaA, iuc*), *Vibrio cholerae* (*hutA, ctx*), and *E. coli* O157:H7 (*shuA, eae, jct*) were prepared (between three and four genomic probes for each amplicon were made). Six unique oligonucleotide probes were designed and prepared for each of the two amplicon sequences (N & M) of the SARS-CoV, while each of these probes were covalently coupled to a unique bead. Finally, after identifying specific genomic regions within RVFV that code for the NSs and L proteins, both amplicons and six unique oligonucleotide probes were designed and prepared. It should be noted that all probes made for this work were generated from genomic sequences that are found in the public domain at:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide&itool=toolbar.

2. TECHNOLOGY OF MULTIPLEXED ASSAYS USING THE LUMINEX xMAP SYSTEM

The Luminex xMAP system provides a rapid and cost effective method to simultaneously detect the presence of specific DNA or RNA sequences of many different bacterial and viral pathogens. The xMAP system utilizes polystyrene microspheres internally dyed with two spectrally distinct fluorochromes. Using precise ratios of these fluorochromes, a suspension array is created consisting of 100 spectrally distinct microsphere sets. Each microsphere set can possess a different genomic probe on its surface, and up to 100 different analytes/sequences can be detected simultaneously in a single reaction vessel. A third fluorochrome, in this case a biotinylated amplicon, acts as a reporter molecule allowing quantification of the biomolecular interaction that has occurred at the microsphere surface. The multiplexing capability of the xMAP system provides an assay detection platform that does not require users to prejudicially screen for specific organisms, but allows a more thorough screen for the presence of biowarfare, food-borne, and environmental pathogens.

3. DEVELOPMENT OF LUMINEX xMAP SCREENING OF AMPLICONS OF ENTEROPATHOGENIC BACTERIA, SARS-COV AND RVFV

Each oligonucleotide probe used in this study was covalently coupled to a unique bead. Amplicons were screened under hybridization conditions (48°C) optimized to permit hybridization of all the probes in the set. The sensitivity of the Luminex xMAP system has been examined in two ways: 1) by spiking gloves with *Shigella dysenteriae* and 2) by calculating the theoretical number of organisms that can be detected after the Polymerase Chain Reaction (PCR). Individual gloves were contaminated with $1.5 \times 10^8$, $1.5 \times 10^6$, or $1.5 \times 10^2$ cfu of *Shigella dysenteriae*. Gels containing the products of the PCR reactions from the gloves were run and clear bands were detected from the 150 cfu sample. The Luminex system has also been shown capable of detecting 1 attamole of DNA per sample ($10^6$ molecules of DNA). After 35 PCR cycles, DNA template from very few bacteria (less than 100) will yield a clearly detectable signal at the attamole level.

## RESULTS

1. BACTERIAL PATHOGENICITY ISLANDS AND ASSOCIATED TOXINS

A number of genes in enteric bacteria are found in pathogens but are generally absent from the non-pathogenic strains. These include iron transport genes such as the heme transport genes (*shu* or *chu*) of *Shigella* and pathogenic *E. coli*, as well as the aerobactin (*iuc*) iron transport system (8-13). The distribution of these genes among a variety of pathogenic and non-pathogenic enteric bacteria has been characterized. These data, along with an analysis of the genome data for other virulence factors, have revealed patterns of genes that are associated with specific pathogens. Many of these virulence genes map within pathogenicity islands found in bacterial genomes, bacteriophages, or plasmids. Table 1 shows a sample of these data.

**TABLE 1. - Virulence Genes**

| Strain | *iuc* | *shu* | *ipaB* | *stx* | *eae* | *sit* | *iro* | *ybt* |
|---|---|---|---|---|---|---|---|---|
| ***Salmonella enterica*** | - | + | + | - | - | + | + | - |
| ***E. coli* K12** | - | - | - | - | - | - | - | - |
| ***Shigella flexneri*** | + | - | + | - | - | + | - | - |
| ***Shigella dysenteriae*** | - | + | + | + | - | + | + | - |
| ***Shigella sonnei*** | - | - | + | - | - | + | - | - |
| ***E. coli* 0157:H7** | - | + | - | + | + | - | - | + |
| **Enteropathogenic *E. coli*** | - | + | - | - | + | - | - | - |
| **Uropathogenic *E. coli*** | + | + | - | - | - | + | + | + |

Each pathotype has a distinctive pattern of virulence genes or accessory genes present in the genome. By identifying the pattern, the species or strain of enteric pathogen can be accurately determined. Because some of these genes are responsible for the symptoms and pathology of the diseases, detection of the gene pattern allows prediction of the type of disease that will result from infection with the strains. For example, identification of a strain that carries genes for the *shiga* toxin (*stx*), attaching and effacing lesions (*eae*) and heme transport (*shu*) indicates that it has the potential to cause hemorrhagic colitis (*E. coli* O157:H7, for example). Because the Radix BioSolutions technology allows us to screen for many different DNA sequences simultaneously, probes for highly conserved regions in each gene were used to detect the presence of the genes, as well as probes to variable regions to identify variants (e.g., *stx1* and *stx2*, or *S. enterica iro* vs. *S. dysenteriae iro*).

Many of these genes map within pathogenicity islands. The virulence genes within the islands, such as the *iuc* genes, may be highly conserved. Other genes or sequences within the island may be unique to that strain. The *iuc* genes in *S. flexneri* and *S. boydii* are essentially identical but other island sequences adjacent to the *iuc* genes are distinct in each species. The location of the island is different in these species. The *iuc* island maps downstream of *selC* in *S. flexneri* but is downstream of *pheU* in *S. boydii*. Probes that recognize the non-*iuc* island sequences and the junctions between the island and the backbone genome can easily distinguish between these two species.

## 2. PROBE SPECIFICITY FOR VIRAL PATHOGENIC FACTOR AMPLICONS

Synthetic DNA probes were designed to detect potential amplicons of specific genomic regions found in pathogenic viruses. In Figure 1a, a negative control (-) shows that no amplification of the PCR products was detected in the absence of probe sets. However, when probes were included in the PCR mixture (+), a homogeneous band at 300 base pairs (bp) was present. Therefore, there are specific amplicons for the N and M genes of SARS-CoV. Likewise, a similar result was obtained for the NSs and L genes for RVFV (Figure 1b.). This gel reveals the homogeneity of the products, therefore concluding there are amplicons for NSs and L genes of RVFV.

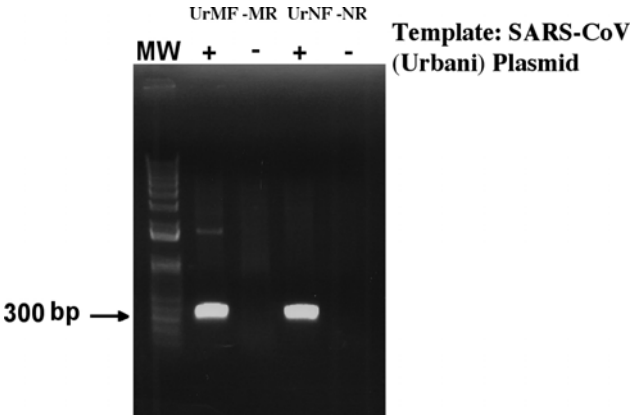


Figure 1a. SARS-CoV M and N PCR amplification products.

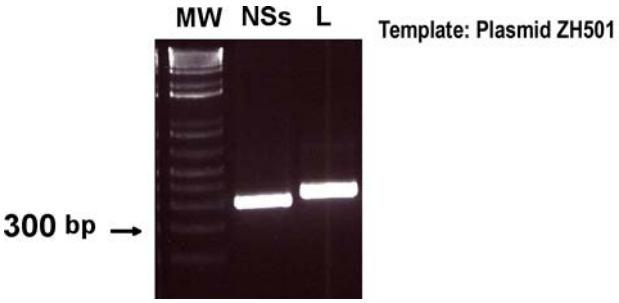


Figure 1b. RVFV NSs and L PCR amplification products.

3. LUMINEX xMAP SCREENING OF AMPLICONS OF ENTEROPATHOGENIC BACTERIA, SARS-CoV AND RVFV

The following two subsections describe the actual application of the specific probes and screening of amplicons using the Luminex xMAP system. These results form the basis for the use of these probes to detect amplicons of threat agents using a platform that will assay several hundred samples per hour.

a. *Shigella dysenteriae*, S. *flexneri*, *Vibrio cholerae*, and *E. coli* O157:H7. The designed probe sets can specifically, and with no apparent false negatives/positives, detect eight enteric pathogenicity islands (Figure 2). Furthermore, Figure 2 illustrates the high degree of specificity and ability of probes to clearly recognize amplicons from specific pathogenicity islands (*hutA*, *eae*, *ipaA*, *stx*, *ctx*, *shuA*, *jct*, and *iuc*). Each bar represents a unique bead with a unique oligonucleotide probe. For seven of the eight amplicon sequences, four probe bead sets were used, and for *hutA* three probe bead sets were used. When the total of 31 probe bound bead sets were used to screen for the 8 individual PCR amplicons, all specific probes were recognized as positives.

b. SARS-CoV and RVFV. Six unique oligonucleotide probes were designed for each of the two amplicon sequences of SARS-CoV. Six unique oligonucleotide probes were designed for each of the two amplicon sequences of RVFV. Each of the probes was covalently coupled to a unique bead. The SARS-CoV M amplicon yielded three high signal intensities relative to background non-binding probes, while the SARS-CoV N amplicon yielded five highly specific signal intensities (Figure 3). With all of the RVFV PCR products, all six probes showed high specificity and signal intensity (Figure 3).

c. False negatives and false positives. Multiple probes for the same amplicon reduce false positive and false negative detection events. This assertion is validated by examination of Figure 2. In this figure, two observations can be made: the first is that probes that were not designed for the amplicons all yielded a similar low signal. The second is that only the appropriate probes for the amplicons reacted strongly. In Figure 2, only the four probes for amplicons *eae*, *ipaA*, *stx*, *ctx*, *shuA*, *jct*, and *iuc* reacted with the appropriate amplicon. *HutA* reacted with three probes. Accordingly, as seen in Figure 3, similar specificities were observed. All of these observations exemplify the reduced likelihood of false negatives and of false positives; however, the real test of false positives requires testing every known strain of organism in every known matrix and this is impractical.

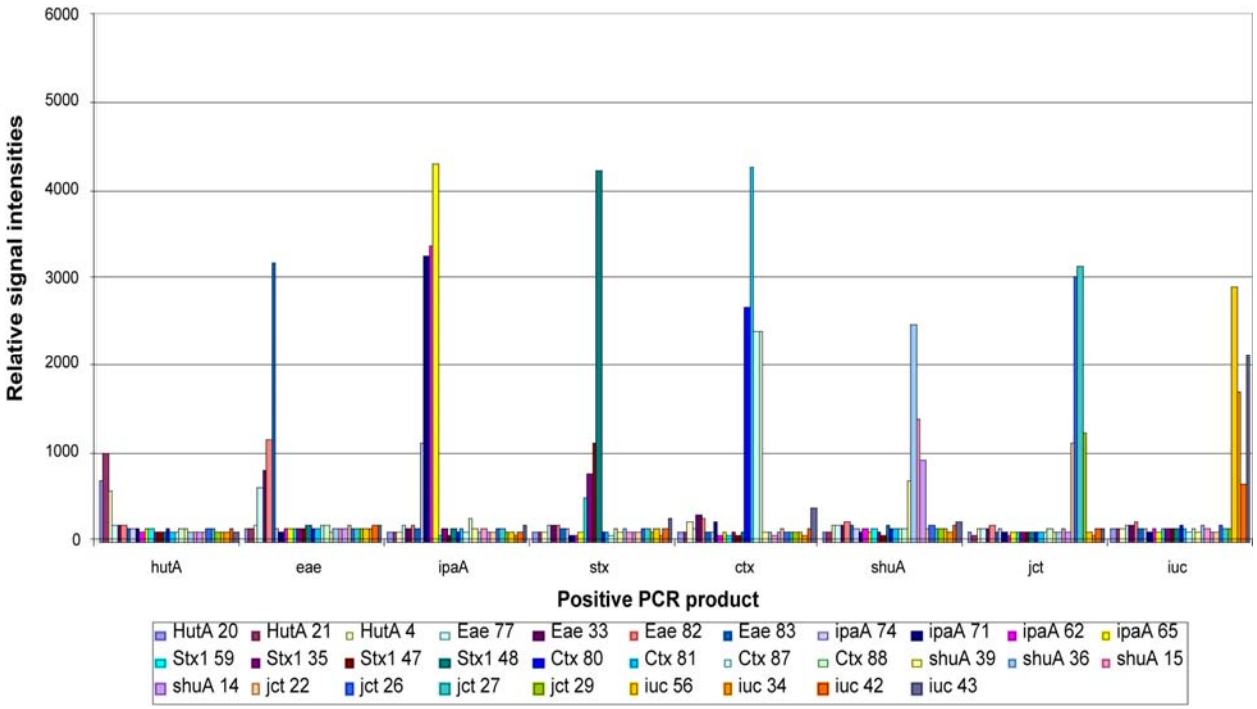


Figure 2. Hybridization of pathogenicity island amplicons to specific polynucleotide probes.

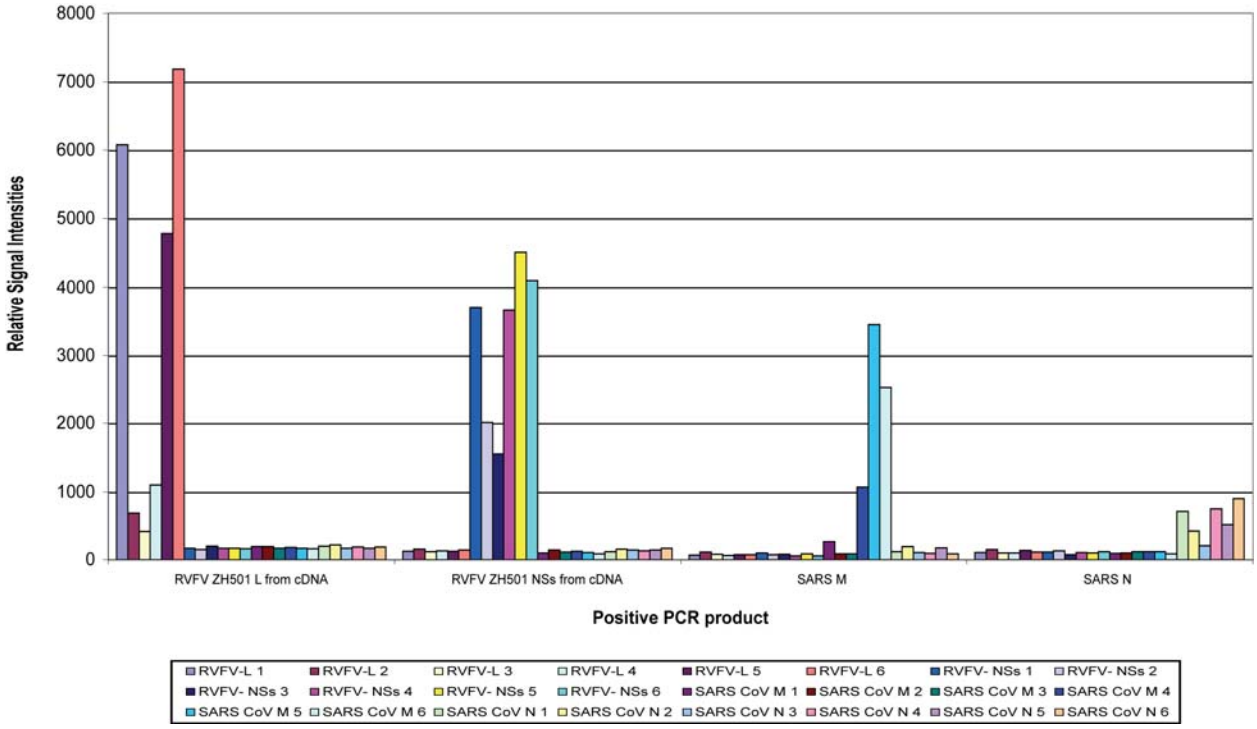


Figure 3. Hybridization of RVFV amplicons (NSs and L) and SARS-CoV amplicons (M and N) to specific polynucleotide probes.

## CONCLUSIONS

Probe sets have been designed and developed to specifically generate PCR products (amplicons) from genes which are characteristic of eight pathogenicity islands in enteropathogenic bacteria. Probe sets for SARS-CoV and RVFV genes associated with pathogenicity have also been prepared. Additionally, all of these pathogenic genomes have been subject to testing using the multiplexed assays applied by the Luminex xMAP System. Importantly, the PCR results provide a proof-of-concept for the development of the PCR probe sets which can be used in a multi-PCR assay kit for the generation of amplicons of interest in a mixture. Additionally, genomic probes can specifically, and with no false negatives, detect the eight enteric pathogenicity islands and the SARS-CoV and RVFV sequences described in this report.

The proposed screening system has commercial applications in office health monitoring, hospital infectious disease detection, environmental, agricultural and food-borne pathogen monitoring. As new pathogenicity islands and factors are discovered, or mutations are determined, new microsphere probe sets that carry new sequences can be included to seamlessly update any developed kit.

## ACKNOWLEDGEMENTS

The research reported in this manuscript was performed in connection with the Countermeasures to Biological and Chemical Threats Program (ECBC, DAAD13-02-C-0079) and with the Center for Strategic and Innovative Technologies (NIH,1-U01-A10533857) of The University of Texas at Austin. The authors would like to acknowledge Tania Tost for her work preparing the manuscript and Cheryl Rae for expert assistance with figure preparation.

## REFERENCES

1. Vennema, H.G., Godeke, G.J., Rossen, J.W., Voorhout, W.F., Horzinek, M.C., Opstelten D.J., and Rottier, P.J., 1996: Nucleocapsid-Independent Assembly of coronavirus-like particles by co-expression of viral envelope protein genes, *EMBO J.* **15**, 2020-2028.

2. de Haan, C.A.M., Smeets, M. Vernooij F, Vennema H., and Rottier P.J.M., 1999: Mapping of the Coronavirus Membrane Protein Domains Involved in Interaction with the Spike Protein, *J. Virol.* **73**, 7441-7452.

3. Narayanan, K., Maeda, A., Maeda, J., and Makino, S., 2000: Characterization of the Coronavirus M protein and Nucleocapsid Interaction in Infected cells, *J. Virol.* **74**, 8127-8134.

4. Narayanan, K. and Makino, S., 2001: Cooperation of an RNA Packaging Signal and a Viral Envelope Protein in Coronavirus RNA Packaging, *J. Virol.* **75**, 9059-9067.

5. Narayanan K., Kim K.H., and Makino S., 2003: Characterization of N Protein Self-Association in Coronavirus Ribonucleoprotein Complexes, *Virus Res.* **2**, 131-140.

6. Lopez, N., Muller, R., Prehaud, C., Bouloy, M. 1995: The L protein of Rift Valley fever virus can rescue viral ribonucleoproteins and transcribe synthetic genome-like RNA molecules, *J. Virol.* **7**, 3972-3979.

7. Billecocq, A., Spiegel, M., Vialat, P., Kohl, A., Weber, F., Bouloy, M., and Haller, O. 2004: NSs Protein of Rift Valley Fever Virus Blocks Interferon Production by Inhibiting Host Gene Transcription, *J. Virol.* **18**, 9798-806.

8. Mills, M. and Payne S.M., 1995: Genetics and Regulation of Heme Iron Transport in *Shigella dysenteriae* and Detection of an Analogous System in *Escherichia coli* O157:H7, *J. Bacteriol.* **11**, 3004-3009.

9. Mills M. and Payne, S.M., 1997: Identification of *shuA*, the Gene Encoding the Heme Receptor of *Shigella dysenteriae*, and Analysis of Invasion and Intracellular Multiplication of a *shuA* mutant, *Infect Immun.* **12**, 5358-63.

10. Purdy, G. and Payne, S.M., 2001: The SHI-3 Iron Transport Island of *Shigella boydii* 0-1392 Encodes the Genes for Aerobactin Synthesis and Transport, *J. Bacteriol.* **183**, 4176-4182.

11. Torres, A and Payne, S.M., 1997: Haem Iron Transport System in Enterohaemorrhagic *Escherichia coli* O157:H7, *Mol. Microbiol.* **23**, 825-833.

12. Vokes, S.A., Reeves, S., Torres, A.G., and Payne, S.M. 1999: The Aerobactin Iron Transport Genes in *Shigella flexneri* are Present within a Pathogenicity Island, *Molecular Microbiology* **1**, 63-73.

13. Wyckoff, E., Duncan, D., Torres, A., Maase, K., and Payne, S.M., 1998: Structure of the *Shigella dysenteriae* Haem Transport Locus and its Phylogenetic Distribution in Enteric Bacteria, *Mol. Microbiol.* **28,** 1139-1152.